Published online 7 February 2013 *Nucleic Acids Research, 2013, Vol. 41, No. 6* ***3805–3818***
doi:10.1093/nar/gkt062

Comprehensive *in vivo* RNA-binding site analyses reveal a role of Prp8 in spliceosomal assembly

Xueni Li, Wenzheng Zhang, Tao Xu, Jolene Ramsey, Lingdi Zhang, Ryan Hill, Kirk C. Hansen, Jay R. Hesselberth and Rui Zhao*

Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, CO 80045, USA

Received November 14, 2012; Revised January 10, 2012; Accepted January 14, 2013

ABSTRACT

Prp8 stands out among hundreds of splicing factors as a protein that is intimately involved in spliceosomal activation and the catalytic reaction. Here, we present the first comprehensive *in vivo* RNA footprints for Prp8 in budding yeast obtained using CLIP (cross-linking and immunoprecipitation)/CRAC (cross-linking and analyses of cDNAs) and next-generation DNA sequencing. These footprints encompass known direct Prp8-binding sites on U5, U6 snRNA and intron-containing pre-mRNAs identified using site-directed cross-linking with *in vitro* assembled small nuclear ribonucleoproteins (snRNPs) or spliceosome. Furthermore, our results revealed novel Prp8-binding sites on U1 and U2 snRNAs. We demonstrate that Prp8 directly cross-links with U2, U5 and U6 snRNAs and pre-mRNA in purified activated spliceosomes, placing Prp8 in position to bring the components of the active site together. In addition, disruption of the Prp8 and U1 snRNA interaction reduces tri-snRNP level in the spliceosome, suggesting a previously unknown role of Prp8 in spliceosomal assembly through its interaction with U1 snRNA.

INTRODUCTION

Pre-mRNA splicing is essential for gene expression in all eukaryotes. Introns are removed through two transesterification reactions (1). In the first reaction, the 2′ hydroxyl group of the branchpoint adenosine residue in the branchpoint sequence (BPS) attacks the phosphate group at the 5′ ss, generating a lariat intermediate. In the second transesterification reaction, the newly released 3′ hydroxyl group from the cleaved 5′ exon attacks the phosphate group at the 3′ ss, releasing the lariat-intron and ligating the two exons. The splicing reaction is catalysed by the spliceosome, a large ribonucleoprotein complex. The spliceosome contains five small nuclear RNAs (U1, U2, U4, U5 and U6 snRNAs) that form five small nuclear ribonucleoproteins (snRNPs) with their associated proteins, plus numerous other protein factors (2). Spliceosome components typically assemble on pre-mRNA in a stepwise manner. Formation of the E complex involves the initial recognition of intron elements by various spliceosomal components. Subsequently, the U2 snRNP joins the spliceosome, interacting with the BPS and forming the A complex. This is followed by the joining of the U4/U6.U5 tri-snRNP and the formation of the B complex, although it is unclear what recruits the tri-snRNP to the spliceosome. Extensive structural rearrangements occur at this stage to form the catalytically active B* complex, which is ready for first step catalytic reaction. During the activation process, the base pairing between the 5′ ss and U1 snRNA as well as the interaction between U4 and U6 are disrupted, and U1 and U4 leave the spliceosome. After the first step has been completed, the spliceosome repositions the substrate for the second catalytic step to form the C complex. The second step is followed by post-catalytic rearrangements to liberate the mature mRNA for export, release the lariat intron to be degraded and recycle the snRNPs.

Prp8 stands out among hundreds of splicing factors as a protein that lies at the heart of the spliceosome [reviewed in (3)]. Prp8 is a U5 and tri-snRNP protein and is also present in the spliceosome. Prp8 is one of the largest (>2000 amino acids in length) and most conserved (human and yeast Prp8 share 61% sequence identity)

*To whom correspondence should be addressed. Tel: +1 303 724 3269; Fax: +1 303 724 3215; Email: rui.zhao@ucdenver.edu
Present addresses:
Wenzheng Zhang, The Samuel Roberts Noble Foundation, Ardmore, OK 73401, USA.
Tao Xu, Department of Microbiology, University of Colorado School of Medicine, Aurora, CO 80045, USA.
Jolene Ramsey, Department of Biology, Indiana University, Bloomington, IN 47405, USA.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

proteins in the nucleus. However, Prp8 has remarkably low-sequence similarity with other proteins, making it difficult to deduce its function from sequence analyses. Structural studies revealed that the C-terminal region of Prp8 contains an MPN (Mpr1, Pad1 N-terminal) domain and an RNase H domain (4–8), but the structure and function of the majority of Prp8 remain unknown. Genetic analysis has identified numerous *prp8* mutants that suppress or exacerbate mutations in pre-mRNA substrates, or other splicing factors known to play a role in spliceosomal activation. For example, mutations in Prp8 have been identified to suppress or enhance splicing defects caused by 5′ ss, 3′ ss or BPS mutations and by mutations in spliceosomal factors that act at the second step of splicing [reviewed in (3)]. Over 40 Prp8 mutations were found to suppress the cold-sensitive phenotype of the *U4-cs1* mutation, which inhibits U4/U6 unwinding, an important step in spliceosomal activation (9). Furthermore, Prp8 mutants suppress mutant alleles of the RNA helicases Brr2 and Prp28 (10). Brr2 and Prp28 are responsible for the unwinding of U4/U6 and U1/5′ ss, respectively, two essential events required for spliceosomal activation. These data led to the hypothesis that Prp8 may be a master regulator of spliceosomal activation.

Ultraviolet (UV) cross-linking experiments have placed Prp8 physically near the spliceosomal catalytic core [reviewed in (3)]. Prp8 is the only spliceosomal protein that extensively cross-links with all three pre-mRNA regions required for splicing (the 5′ ss, the 3′ ss and the BPS) as well as with the U5 and U6 snRNAs. These observations led to the speculation that Prp8 may help form or stabilize the active site or contribute functional groups to the catalytic reaction. Although these cross-linking experiments have provided important information on the function of Prp8 in splicing, they rely on the incorporation of a photoactivatable nucleotide analogue (e.g. 4-thiouridine) at a specific position of radiolabelled snRNAs or pre-mRNA, assembly of these RNAs into snRNPs or spliceosomes, *in vitro* UV cross-linking, followed by immunoprecipitation to detect cross-linked proteins. Consequently, only a small number of nucleotide positions have been investigated in these cross-linking studies. For example, only one nucleotide position in U6, the snRNA thought to be most likely involved in the catalytic reaction, was examined (11).

CLIP (cross-linking and immunoprecipitation) (12) or CRAC (cross-linking and analyses of cDNAs) (13) experiments followed by high-throughput sequencing is an ideal approach to comprehensively identify the *in vivo* RNA-binding sites of Prp8. In CLIP experiments, intact cells were exposed to 254-nm UV light, which specifically cross-links protein and RNAs under physiological conditions. It takes advantage of the natural reactivity of the nucleic acid bases and specific amino acids, such as Cys, Lys, Phe, Trp and Tyr, at 254-nm UV light (14,15) and avoids the indirect cross-linking found with other reagents, such as formaldehyde. After limited RNase treatment, the protein–RNA complex is immunoprecipitated or pulled down using an affinity tag. Covalent protein–RNA cross-links formed by UV irradiation are used to stringently purify specific protein–RNA complexes using sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS–PAGE) separation. After proteinase K digestion to remove bound proteins, RNAs are amplified through reverse transcriptase–polymerase chain reaction (RT–PCR), analysed by high-throughput sequencing and mapped to the genome to identify bound RNA fragments. In the CRAC method, an additional step consisting of nickel resin purification of the protein–RNA complex under denaturing conditions is added before SDS–PAGE to further remove any contaminating proteins and RNAs (13).

Here, we report the comprehensive *in vivo* RNA-binding sites of yeast Prp8 identified using CLIP/CRAC. The majority of reads map to U5 and other snRNAs, providing *in vivo* footprints of Prp8 on these snRNAs. These footprints contain previously known Prp8 cross-linking sites on the U5 and U6 snRNAs and pre-mRNAs, and they revealed novel sites of cross-linking on U1 and U2 snRNAs. We demonstrate that Prp8 directly cross-links with U2, U5 and U6 snRNAs, as well as pre-mRNA, in purified activated spliceosomes, placing Prp8 in a position to bring all components of the active site together. In addition, disruption of the Prp8 and U1 snRNA interaction reduces tri-snRNP level in the spliceosome, suggesting a new role of Prp8 in spliceosomal assembly through its interaction with U1 snRNA.

MATERIALS AND METHODS

Yeast strains and plasmids

CLIP and CRAC experiments were performed in yeast strains yJU75 [*MATa*, *ade2 cup1Δ::ura3 his3 leu2 lys2 prp8Δ::LYS2 trp1*; *pJU169* (*PRP8 URA3 CEN ARS*)] (16) (gift of C. Guthrie) carrying pRS413/GPD (Glycerol-3-Phosphate Dehydrogenase)-Prp8-TAP (Tandem Affinity Purification) or pRS413/GPD-Prp8-HTP (WT (Wild Type) Prp8 under a GPD promoter) and the yeast TAP collection strain with endogenous chromosomal Prp8 TAP tagged (17).

To evaluate the interaction between Prp8 and U1 snRNA as well as the formation of U1 snRNP, we generated HTP-tagged endogenous Prp8 or U1-70K strain using PCR-based genomic integration (18) in yeast strain yJU46 [*MATα*, *his3*, *trp1*, *lys2*, *ade2*, *snr19::LYS2*, (*U1 WT*, *URA3 CEN ARS*)] (gift of C. Guthrie). The U1 Δ184–312 plasmid was generated from pSE538 Snr19 (gift of C. Guthrie) using QuikChange mutagenesis (Stratagene). These plasmids were transformed into the HTP-tagged Prp8 or U1-70K strain, and wild-type U1 snRNA was replaced by plasmid shuffling.

CLIP and CRAC experiments

Details of the CLIP and CRAC experiments are described in the Supplementary Data. Raw sequencing reads containing 5′ barcodes were de-multiplexed before analyses. Sequencing reads were mapped to the *S. cerevisiae* genome (version sacCer1), the snRNAs and pre-mRNAs (19,20) using cross_match (http://www.phrap.org/phredphrapconsed.html). Deletions within sequencing reads were identified from cross_match alignments and correlated

with snRNA sequences and annotated splice sites in pre-mRNAs (19,20).

Gel shift experiment

We performed gel shift experiments on RNAs extracted from cross-linked Prp8–RNA sample to examine the identity of these RNAs. These RNAs were generated the same way as in a typical CLIP/CRAC experiment but without linker ligation. The RNAs were incubated separately with no oligo, or oligos that are reverse complimentary to U5 (15–33), (50–72), (67–90), (112–133), (158–180) at 37°C for 15 min, analysed on a 9% native acrylamide gel followed by autoradiograph.

Real-time PCR

RNAs extracted from purified Prp8–RNA complexes (derived from whole-cell extract or purified B and B^{act} complexes) were reverse transcribed using random hexamers and analysed by real-time PCR. Primers used for real-time PCR (sequences shown in Supplementary Data) are selected to be on either the 5′- or 3′-end of major cross-linking sites observed in our CLIP/CRAC experiments so that these cross-linking sites will not interfere with the PCR reaction.

We also used real-time PCR to evaluate the splicing phenotype of several genes (*TUB1*, *ACT1*, *RPL21a* and *RPL17b*) in WT and U1 Δ184–312 strains. Total RNAs were extracted from each strain, reverse transcribed using the Random Primer Mix (New England Biolabs) and quantified with real-time PCR using a set of primers specific for the intron to evaluate the level of intron-containing pre-mRNAs and another set of primers specific for the exon to evaluate the level of total mRNAs. Primers for *ACT1*, *RPL21a* and *RPL17b* are reported in a study by Pliess *et al.* (21), and primers for *TUB1* are listed in the Supplementary Data.

Purification of U5 snRNP, tri-snRNP and spliceosomal B and B^{act} complexes

U5 snRNP and tri-snRNP were purified essentially as described previously (22). The spliceosomal B and B^{act} complexes were assembled and purified as described previously (23). Details of the assembly and purification procedure are described in the Supplementary Data.

Evaluation of U1 snRNP formation

U1 snRNP was affinity purified by pulling down HTP-tagged U1-70K from yeast cell extract using IgG resin and eluted by Tobacco Etch Virus (TEV) protease cleavage. Proteinase K was used to digest away proteins in purified U1 snRNP, and U1 snRNA was quantified using in solution hybridization with a U1-specific primer (24). The purified U1 snRNP was tube gel digested with trypsin (25) and subjected to LC–MS (Liquid Chromatography-Mass Spectrometry)/MS analysis (26).

RESULTS

CLIP and CRAC experiments

To identify the *in vivo* RNA-binding sites of Prp8, we carried out CLIP experiments using yJU75 (16), which has its endogenous *PRP8* deleted and carries a plasmid with C-terminal TAP-tagged *PRP8* under an overexpressing GPD promoter, and the yeast TAP collection strain with chromosomal *PRP8* TAP-tagged at the C-terminus (17). Figure 1a shows that the UV-treated sample demonstrated an obvious band whose size is slightly larger than Prp8 protein alone after SDS–PAGE and autoradiograph, whereas the non–UV-treated sample does not. The CLIP results from the two different strains are essentially identical (Table 1). Taking the yJU75 strain overexpressing *PRP8-TAP* as an example, Illumina sequencing produced ~2.5 millions reads that are mapped to the yeast genome. The majority of reads (75%, ~1.9 million) map onto one of the snRNAs (Table 1). There are also significant reads (1.3%, ~32 000) that map to intron-containing genes. Of the remaining reads, most (18%, ~45 700) map to ribosomal RNA, which is a common contaminant in CLIP experiments because of their abundance. We also performed a control CLIP experiment using the yJU75 strain expressing untagged Prp8. The majority of reads (87%) map to rRNA, and only ~1% of the total reads map onto snRNA (Table 1), validating that the RNA-binding sites of Prp8 we identified in CLIP are highly specific.

To further rule out possible contaminating RNAs associated with Prp8 in the CLIP procedure, we carried out a CRAC experiment developed by the Tollervey laboratory (13). We replaced the TAP tag with the HTP (9× His–TEV protease site–Protein A) tag and subjected the Prp8–RNA sample to an additional stringent denaturing purification using nickel resin under 6 M guanidine–HCl. Overall, these sequencing results are similar to the CLIP experiments, except for the loss of the U4 snRNA-binding sites. We will mostly focus on the CRAC results in the following discussions, but we will compare the CRAC and CLIP results on U4 snRNA.

Prp8 predominantly binds to U5 snRNA

Our initial CLIP/CRAC experiments show that U5 snRNA is robustly represented in sequencing libraries (with ~50% of the total reads mapped to U5 snRNA) and is the most significant binding site of Prp8. However, several observations in our initial CLIP/CRAC experiments led us to modify the original CLIP/CRAC protocol to identify the binding sites of Prp8 on U5 snRNA more precisely. In a typical CLIP/CRAC experiment, RNA fragments between 20 and 50 nt are converted to cDNA, amplified and sequenced. However, we noticed that the vast majority of RNA fragments associated with Prp8 is much larger at ~75 nt (Figure 1b, first lane showing RNA recovered after proteinase K digestion of Prp8 without RNA linker ligation). Treatment with increased RNase A/T1 dose (increased 25-folds to 1.25 U of RNase A and 50 U of RNase T1 per ml of cell lysate) and RNase V1 (which cleaves

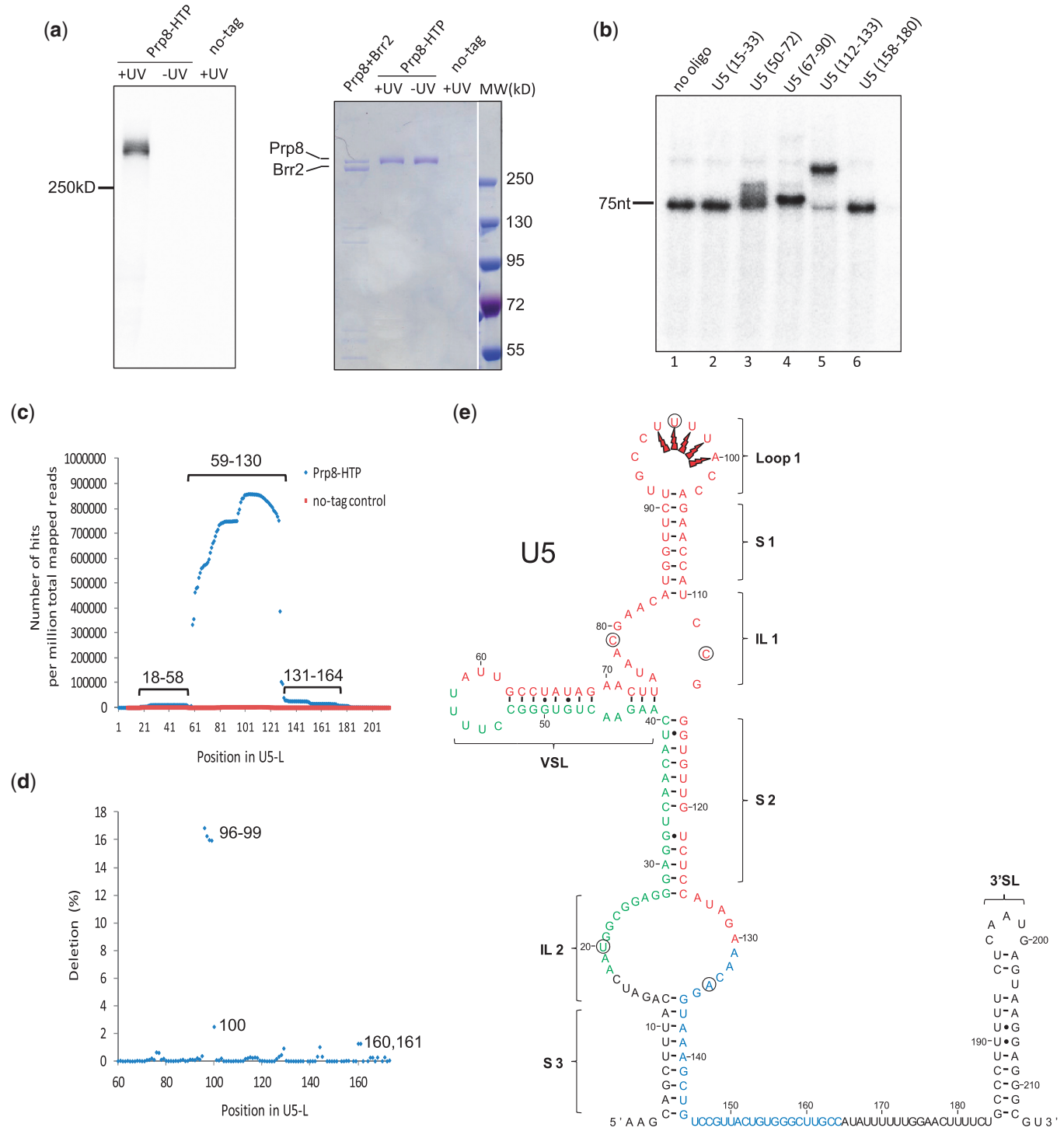


Figure 1. Prp8-binding sites on U5 snRNA. (**a**) Left: A typical autoradiograph in a CRAC experiment after radiolabelling the RNA followed by SDS–PAGE purification. HTP-tagged Prp8 with UV treatment generates an obvious radioactive band corresponding to the Prp8–RNA complex, whereas the controls (no-tag or non-UV) do not. Right: The CRAC samples were stained with Coomassie blue to demonstrate a clean single-Prp8 band in the Prp8–HTP sample (with or without UV treatment). A Prp8 and Brr2 complex was loaded in the first lane as a control to demonstrate positions of the Prp8 and Brr2 proteins. (**b**) The majority of RNAs (5′ ^{32}P-labelled) recovered from cross-linked Prp8–RNA samples in CLIP experiments are ~75 nt in length on a native polyacrylamide gel followed by autoradiograph (lane 1). Oligos complementary to positions 50–72, 67–90 and 112–133 in U5 snRNA can shift this 75-nt band in a gel shift experiment (lanes 3–5), whereas two other oligos (positions 15–33 and 158–180) cannot (lanes 2 and 6). (**c**) Number of reads in a CRAC experiment mapped to different positions in U5 snRNA reveals a major peak between positions 59 and 130 and two smaller peaks around positions 18–58 and 131–175. (**d**) Percentage of reads containing deletions at each position of U5 snRNA reveals a large number of deletions at positions 96–99. (**e**) Major sequencing reads and deletion sites (marked by bolts) mapped to the predicted secondary structure of U5 snRNA (27). Positions with the most abundant sequencing reads are in red, and the two regions with lesser reads are in green and blue. Circles designate cross-linking sites identified in previous site-directed *in vitro* cross-linking experiments (28). S, IL, SL and VSL stand for stem, internal loop, stem loop and variable stem loop, respectively.

Table 1. Summary of sequencing reads in CLIP/CRAC experiments.

	no-tag control		Prp8–TAP endogenous (CLIP)		Prp8–TAP overexpression (CLIP)		Prp8–HTP (CRAC)	
	Hits	Percent total	Hits	Percent total	Hits	Percent total	Hits	Percent total
Mapped reads	1 598 694		1 670 261		2 528 593		630 329	
snRNAs	**16 877**	**1.06**	**1 243 717**	**74.46**	**1 902 448**	**75.24**	**388 845**	**61.69**
rRNA	1 386 253	86.71	218 102	13.06	456 762	18.06	162 932	25.85
tRNAs	151 383	9.47	118 290	7.08	85 399	3.38	41 261	6.55
Intronic genes	**2290**	**0.14**	**25 032**	**1.50**	**32 254**	**1.28**	**13 117**	**2.08**
Others	41 891	2.62	65 120	3.90	51 730	2.05	24 174	3.84

snRNAs and intronic genes (shown in bold) are clearly enriched in sequencing reads of TAP or HTP-tagged Prp8 compared with the no-tag control.

double-stranded RNA) did not alter the size of this band (data not shown). Considering that our initial CLIP/CRAC experiments indicate that the Prp8:U5 snRNA interaction is the most abundantly recovered cross-linking event, we reasoned that this 75-nt RNA is part of U5 snRNA. To map this 75-nt RNA within U5 snRNA, we performed gel shift experiments using oligonucleotide probes complementary to different positions of U5 snRNA. Probes complementary to positions 50–72, 67–90 and 112–133 can shift these RNAs on a native polyacrylamide gel (the different degrees of shift on the gel is likely because of the binding of these probes at different positions of U5), whereas probes complementary to positions 15–33 and 158–180 are unable to undergo a shift (Figure 1b). These gel shift experiments demonstrate that Prp8 binds predominantly on a 75-nt region between positions 50 and 133 of U5 snRNA, protecting it from RNase digestion. We note that the sensitivity of these gel shift experiments is much lower than CLIP/CRAC followed by high-throughput sequencing. It is likely that there are other Prp8 RNA-binding sites in addition to this 75 nt in U5 snRNA that can only be revealed by analysing the high-throughput sequencing data.

Because a canonical CLIP/CRAC procedure typically recovers 20–50 nt RNA fragments and lacks longer U5 snRNA fragments bound by Prp8, we modified the CLIP/CRAC protocol to more precisely define Prp8-binding sites on U5 snRNA (the remaining Prp8-binding sites in the cell are identified using a typical CLIP/CRAC protocol). We treated the cells with a higher RNase dose (5-fold above what is used in a typical CLIP/CRAC experiment) and sequenced cDNA libraries corresponding to 20–80 nt in length using Illumina HiSeq2000 with a 75-nt read length. Comparison of the CRAC reads of HTP-tagged Prp8 (CLIP results are similar) and the no-tag control demonstrates clearly that the major binding site of Prp8 on U5 is between nucleotides 59 and 130 (Figure 1c), consistent with our gel shift experiments. Two additional regions (positions 18–58 and 131–164) have lower numbers of reads but are consistently above background (i.e. no-tag control). We will discuss the validity of these two regions as Prp8-binding sites in conjunction with deletion analyses described later in the text.

We next analysed mutations in sequencing reads to identify direct Prp8:U5 snRNA cross-linking sites. In CLIP/CRAC experiments, a few amino acid residues often remain cross-linked to the RNA targets after proteinase K digestion. These adducts led to errors in the reverse transcription reaction, which subsequently result in mutations in the sequencing reads. Zhang and Darnell compared the deletion, insertion and substitution of sequencing reads of CLIP data of Nova with known Nova-binding sites (29). They found that reverse transcriptase often skips the cross-linked amino acid:RNA adduct, resulting in a cDNA with a deletion at this site. Consequently, deletions—but not substitutions or insertions—in sequencing reads correlate with known Nova-binding sites. We analysed the deletion distribution between the HTP-tagged and no-tag control sequence reads and found that the major cross-linking sites between Prp8 and U5 snRNA are positions 96–99 in the invariant loop 1, with ~15% of the sequence reads carrying a deletion at these nucleotide positions (Figure 1d). Because positions 96–99 are all uridine residues, we cannot unambiguously determine which nucleotide was deleted. To address this, we distributed the total number of reads containing deletions in this region evenly among positions 96–99. It is possible that one or more nucleotides between positions 96 and 99 cross-link with Prp8. In addition to loop 1, 1–2% of reads spanning positions 100 and 160–161 have deletions in our CRAC data sets (Figure 1d). We do not observe significant deletions between positions 18 and 58, suggesting no direct cross-linking between this region and Prp8 in our CRAC experiments. The reason we observe positions 18–58 in our sequencing reads is likely because the extensive base pairing between this region and the rest of U5 snRNA (forming the stem region of both VSL and S2) (Figure 1e) that may not be completely disrupted during the CLIP/CRAC procedure. Kudla *et al.* also noted undisrupted RNA base pairing after nickel purification under 6 M guanidine–HCl in CRAC experiments (30).

Prp8 binds to U6, U1 and U2 snRNAs and intronic pre-mRNAs

U6 snRNA contains the second most abundant sequencing reads, spanning positions 12–86 (Figure 2a). The highest number of reads mapped between positions 44 and 70, forming a plateau in the signal (~61 000 reads per million of total mapped reads) (Figure 2a), and suggesting that positions 44–70 are well protected from

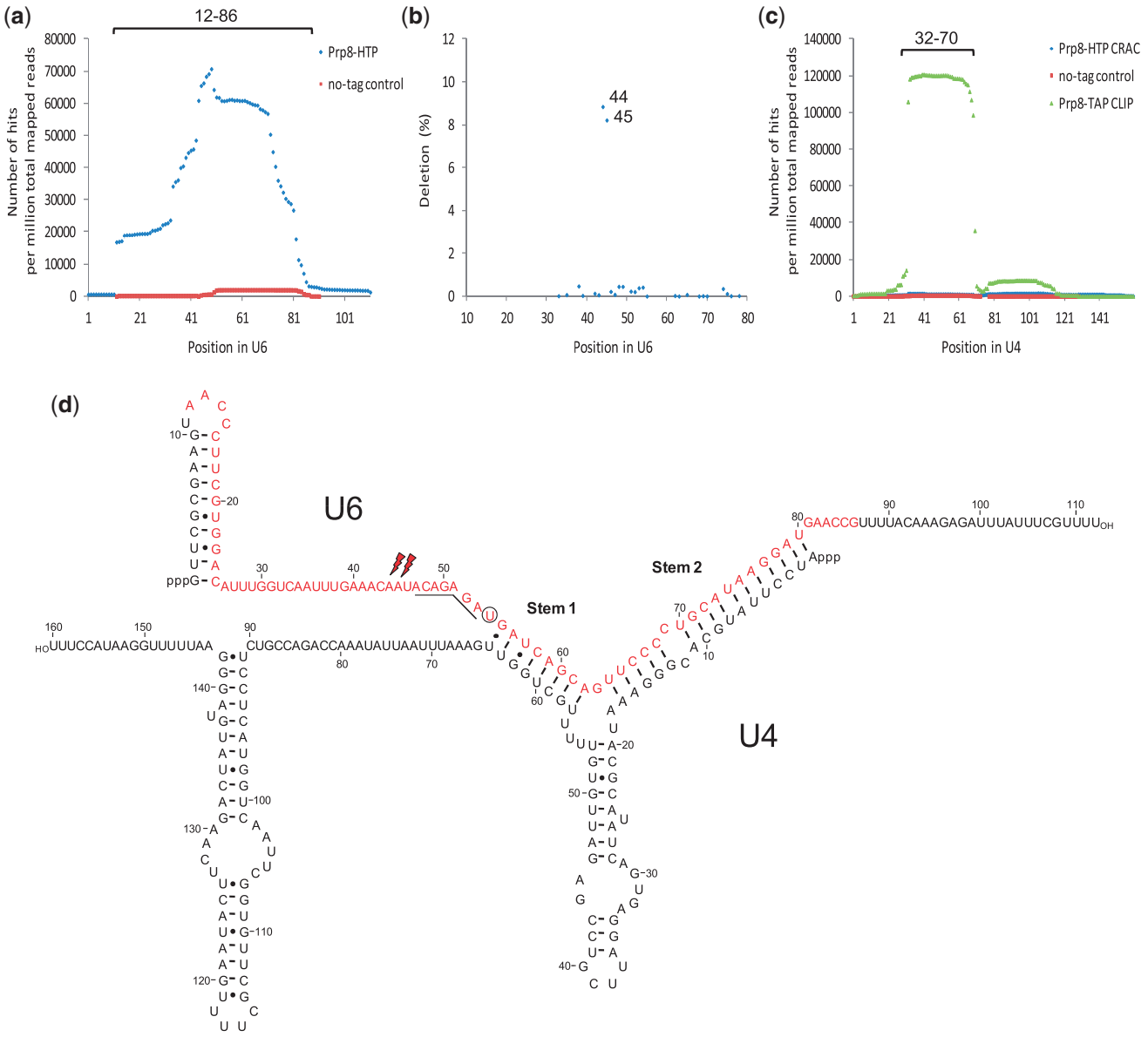


Figure 2. Prp8-binding sites on U6 and U4 snRNAs. (**a**) Number of sequencing reads mapped to different positions in U6 snRNA reveals a major peak between positions 12 and 86 in U6 snRNA. (**b**) Percentage of reads containing deletions at each position of U6 snRNA reveals a large number of deletions at positions 44–45. (**c**) Number of sequencing reads from a CLIP experiment (green) mapped to different positions in U4 snRNA reveals a major peak between 32 and 70 nt in U4 snRNA, but this peak no longer exists in the CRAC experiment (blue). (**d**) Major sequencing reads (red) and deletion sites (marked by lightening bolts) mapped to the predicted secondary structure of U4/U6 snRNA (31). The invariant ACAGAGA box is underlined. The circle designates the cross-linking site identified in previous site-directed *in vitro* cross-linking experiments (11).

RNase digestion by Prp8. Analyses of deletions in sequencing reads demonstrate that the main cross-linking site in U6 is in positions 44–45, with ~8% of the sequencing reads at each of these positions containing deletions (Figure 2b).

The Prp8-binding sites on U4 snRNA exhibited large differences between the CLIP and CRAC experiments. The CLIP sequencing results suggest that Prp8 binds positions 32–70 of U4 snRNA (Figure 2c). However, after a more stringent denaturing purification, the CRAC sequencing results show that Prp8 does not significantly cross-link to this region of U4 snRNA, compared with the no-tag control (Figure 2c). The U4-binding site at positions 32–70 observed in the CLIP experiment could be due to other contaminating proteins or the base pairing of this region of U4 with U6, which is not completely disrupted in the CLIP experiments.

Our CRAC sequencing results demonstrate two major binding sites on U1. The first site is centred at positions 153–182 spanning positions 145–191 (Figure 3a). The second site is centred at positions 292–304 spanning positions 224–329 (Figure 3a). The total number of reads in U1 snRNA is too low to generate reliable deletion analyses results.

Similar analyses demonstrate that U2 snRNA contains a major sequencing reads enrichment area on U2

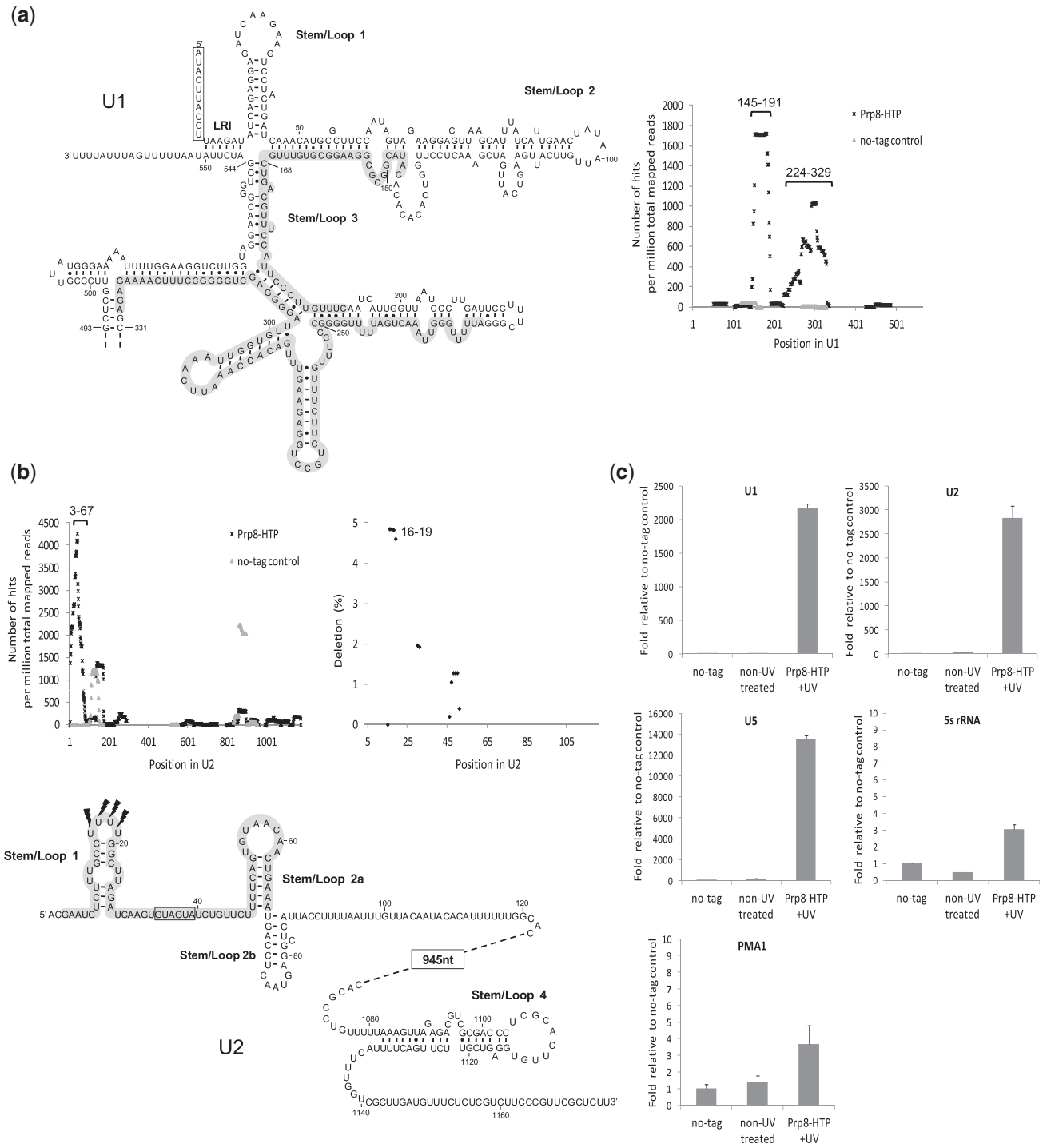


Figure 3. Prp8-binding sites on U1 and U2 snRNAs. (**a**) Number of sequencing reads mapped to different positions in U1 snRNA reveals major peaks at positions 145–191 and 224–329 in U1 snRNA. These positions are also highlighted in grey shade in the proposed secondary structure of U1 [adapted from (32)]. The 5′ ss interacting region in U1 snRNA is boxed. LRI, long range interaction region. (**b**) Number of sequencing reads mapped to different positions in U2 snRNA reveals a major peak at positions 3–67 in U2 snRNA. Percentage of reads containing deletions at each position of U2 snRNA reveals a significant number of deletions at positions 16–19, which is consistently present in all our CRAC data sets. Positions 3–67 are also highlighted in grey shade in the proposed secondary structure of U2 [adapted from (33) and (34)]. Nucleotides that interact with the BPS are boxed. (**c**) Real-time PCR quantification of U1 and U2 snRNAs associated with Prp8 after UV cross-linking. All samples are normalized to the no-tag sample. U5 snRNA is used as a positive control. 5S rRNA and PMA1 mRNA are used as negative controls.

that centres at position 39 spanning positions 3–67 (Figure 3b). There is another peak at position 136, but the same peak is also found in the control and is, therefore, unlikely to be a genuine Prp8-binding site. The major cross-linking site that is consistent in all CRAC data sets spans positions 16–19 (deletion rate ~5% at each position, Figure 3b), suggesting a direct cross-linking site at one or more nucleotides in these positions in U2 snRNA.

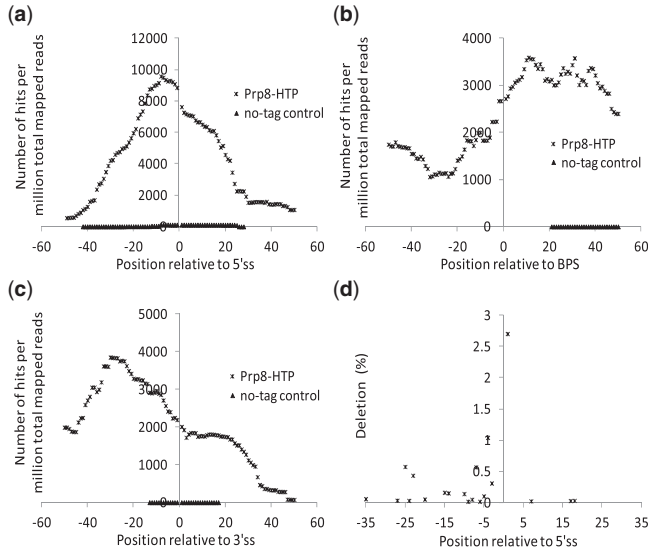


Figure 4. Prp8-binding sites on intron-containing pre-mRNAs. All yeast intron-containing genes are aligned at their 5′ ss, BPS and 3′ ss. A number of sequencing reads are mapped to these intron-containing genes relative to (**a**) the 5′ ss, (**b**) BPS and (**c**) 3′ ss. (**d**) Percentage of reads containing deletions at each position relative to the 5′ ss of intron-containing genes reveals a significant number of deletions at the +1 position.

Because the number of reads mapped to U1 and U2 snRNAs is relatively low, we performed further biochemical experiments to validate whether Prp8 cross-links with U1 and U2 snRNAs. We carried out CRAC experiments but omitted RNase digestion. RNAs recovered were detected using real-time RT–PCR with primers specific for U1 and U2 snRNAs. We also performed real-time PCR experiments using primers specific for U5 snRNA (as a positive control) as well as PMA1 (an abundant mRNA that lacks an intron) and 5S rRNA as negative controls. We showed that U5, U1 and U2 snRNAs are all significantly enriched in the UV cross-linked samples compared with the no-tag and non–UV-treated samples (Figure 3c), whereas the negative controls do not show significant differences among these samples (Figure 3c). These results support our interpretation of the CRAC data that Prp8 directly cross-links with U1 and U2 snRNAs.

In addition to the snRNAs, the other group of RNAs that are significantly represented in our CLIP/CRAC samples compared with the no-tag control is intron-containing pre-mRNAs (Table 1). We aligned the ~300 intron-containing genes in budding yeast (19,20) on their 5′ ss, 3′ ss and BPS. Compared with the no-tag control, there is a clear peak of sequencing reads at the 5′ ss, BPS and 3′ ss (Figure 4a–c). Because the distance between the BPS and 3′ ss is generally short (i.e. between 10 and 50 nt) (19,20), there are overlaps in the signal generated for these two positions. Deletion analyses suggest that the predominant cross-linking site at the 5′ ss is +1 (Figure 4d). The total number of reads around the BPS and 3′ ss is too low to generate reliable deletion analysis results.

Prp8-binding sites in U5 and tri-snRNP

Our aforementioned results define Prp8-binding sites in the entire yeast cell (we will refer to these as the 'whole-cell data set'). We next examined the Prp8-binding sites in specific snRNP complexes: the U5 and tri-snRNPs. We treated yeast cells with UV radiation, then purified U5 and tri-snRNPs using the TAP tag on Prp8. We confirmed our purification of the U5 and tri-snRNP complexes by silver staining of RNAs extracted from each fraction of the glycerol gradient, Coomassie staining of proteins in these fractions and mass spectrometry (Figure 5a). Purified U5 and tri-snRNP were treated with limited RNase, and Prp8–RNA complexes were further purified by Nickel resin under denaturing conditions. Libraries prepared from these samples were sequenced to generate reads corresponding to RNAs cross-linked to Prp8 in U5 and tri-snRNPs.

The binding sites of Prp8 on U5 snRNA are similar in U5 snRNP and tri-snRNP (Figure 5b). In both complexes, the major Prp8 footprint (positions 59–130) is identical to what we observed in the whole-cell data set. In both U5 and tri-snRNP complexes, the secondary binding site at positions 131–164 is more prominent than the whole-cell data. The lower number of sequencing reads in this region (positions 131–164) in the whole-cell data set may be caused by the much higher (>10-fold) RNase dose in the whole-cell CLIP/CRAC experiments compared with what is used to treat purified U5 and tri-snRNPs, which could reduce the amount of RNAs that are less well protected by Prp8 (e.g. positions 131–164 of U5).

As expected, the sequencing reads mapped to U4 and U6 snRNAs are not above background levels in U5 snRNP (Figure 5c and d). In the tri-snRNP, there are mapped reads between positions 76 and 160 on U4 snRNA (Figure 5c). This likely represents a weak Prp8-binding site on U4 snRNA, which is not observed in the whole-cell CRAC data, possibly because of the higher RNase dose used and the much lower abundance of U4 snRNA in the whole-cell CRAC data sets (the tri-snRNP data set greatly enriches U4 and U6 sequence reads). The predominant binding site of Prp8 on U6 snRNA is between positions 12 and 86 (Figure 5d), identical to the Prp8-binding site on U6 in the whole-cell data. There is a secondary binding site of Prp8 on U6 snRNA between positions 87 and 110, which is not observed in the whole-cell data set (possibly because of the higher RNase treatment and the lower percentage of U6 reads in the whole-cell data set).

Prp8–RNA interactions in the spliceosomal B and B^{act} complexes

To examine the Prp8-binding sites in the spliceosome, we assembled and purified the B and B^{act} complex using the yeast M3-Act1 Δ6 pre-mRNA substrate in the presence of 0.05 mM adenosine triphosphate (ATP) (for the B complex) or 2 mM ATP (for the B^{act} complex) (23) (Figure 6a). B^{act} is an activated spliceosomal complex immediately before B* that lacks U1 and U4, but it has not undergone the conformational change catalysed by the DEAH-box helicase Prp2 to become B*, which will

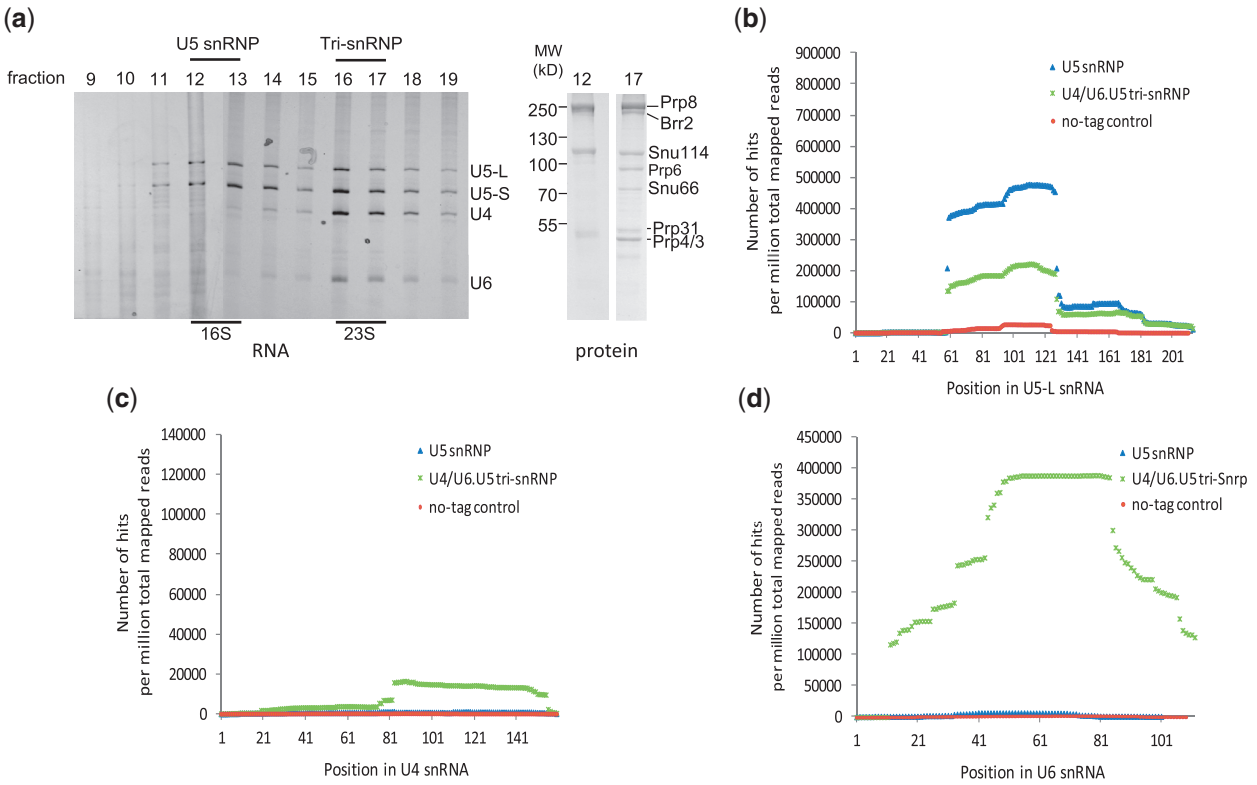


Figure 5. Prp8-binding sites in U5 and tri-snRNPs. (**a**) Purification of U5 and tri-snRNP on glycerol gradient. Positions of bacterial 16S and 23S ribosomal RNAs as sedimentation coefficient standards are also indicated. RNA extracted from each fraction after silver staining and mass spectrometry of selected protein bands confirmed that fractions 12–13 correspond to U5 snRNP, and fractions 16–17 correspond to the tri-snRNP. (**b**–**d**) Prp8 sequencing reads mapped to each position of U5, U4 and U6 snRNAs in U5 snRNP (blue) and tri-snRNP (green).

immediately go through the first step reaction. CRAC experiments using purified B and B^{act} did not generate significant reads compared with the no-tag control, possibly because of the very low quantity of purified B and B^{act} that we obtained. The multiple purification steps in CRAC experiments, as well as incomplete RNase digestion and linker ligation efficiency can all contribute to the further reduction of usable sequencing reads in the sample.

Instead of the regular CRAC experiments, we UV cross-linked the purified B and B^{act} complex, omitted the RNase digestion and linker ligation steps, purified the HTP-tagged Prp8:RNA complex and examined the extracted RNAs using real-time PCR with primers specific for each snRNA as well as the intron region of *act1* and 5S rRNA (as a negative control). U2, U5 and U6 snRNA and *act1* are significantly enriched (~6–60-fold) in the UV cross-linking B^{act} complex compared with the non-UV cross-linking samples, whereas the 5S rRNA control was not significantly enriched (Figure 6b), indicating that Prp8 cross-links with the U2, U5 and U6 snRNAs and *act1* in the B^{act} complex. Prp8 cross-links to U5 and *act1* pre-mRNA at similar levels in the B^{act} and B complexes, but it has a higher level of cross-linking with U2 and U6 snRNA in the B^{act} complex compared to the B complex. Prp8 also has significant cross-linking with U1 and U4 snRNAs in the B complex. These contacts are significantly reduced in the B^{act} complex, but they do not completely disappear, likely because the B^{act} complex assembled and purified under this condition often contains residual amount of U1 and U4 snRNAs (Figure 6a).

Disruption of Prp8 and U1 binding reduces tri-snRNP in spliceosomal assembly

To understand the function of the novel interaction between Prp8 and U1 snRNA, we used a previously reported U1 Δ184–312 construct (32) that contains a partial deletion of the first and second binding sites of Prp8 on U1 snRNA. U1 Δ184–312 grows slightly slower (1.2×) than the WT at 30°C (32) and significantly slower at 16°C (Figure 7a). U1 Δ184–312 exhibits significant pre-mRNA accumulation for several genes that we evaluated using quantitative RT–PCR (Figure 7b), indicating that U1 Δ184–312 causes a splicing defect. U1 Δ184–312 does not affect U1 snRNA levels in the cell (32) (Figure 7c). We further purified U1 snRNP through affinity purification using a HTP tag on the U1-70K protein and demonstrated that U1 snRNP in both the U1 Δ184–312 and WT strains have similar levels of U1 snRNA (by in solution hybridization) and protein components (by mass spectrometry analyses) (Figure 7d), suggesting that U1 Δ184–312 does not affect U1 snRNP formation. We performed the CRAC experiment omitting the RNase treatment step and quantified U1 snRNA cross-linked to Prp8 using

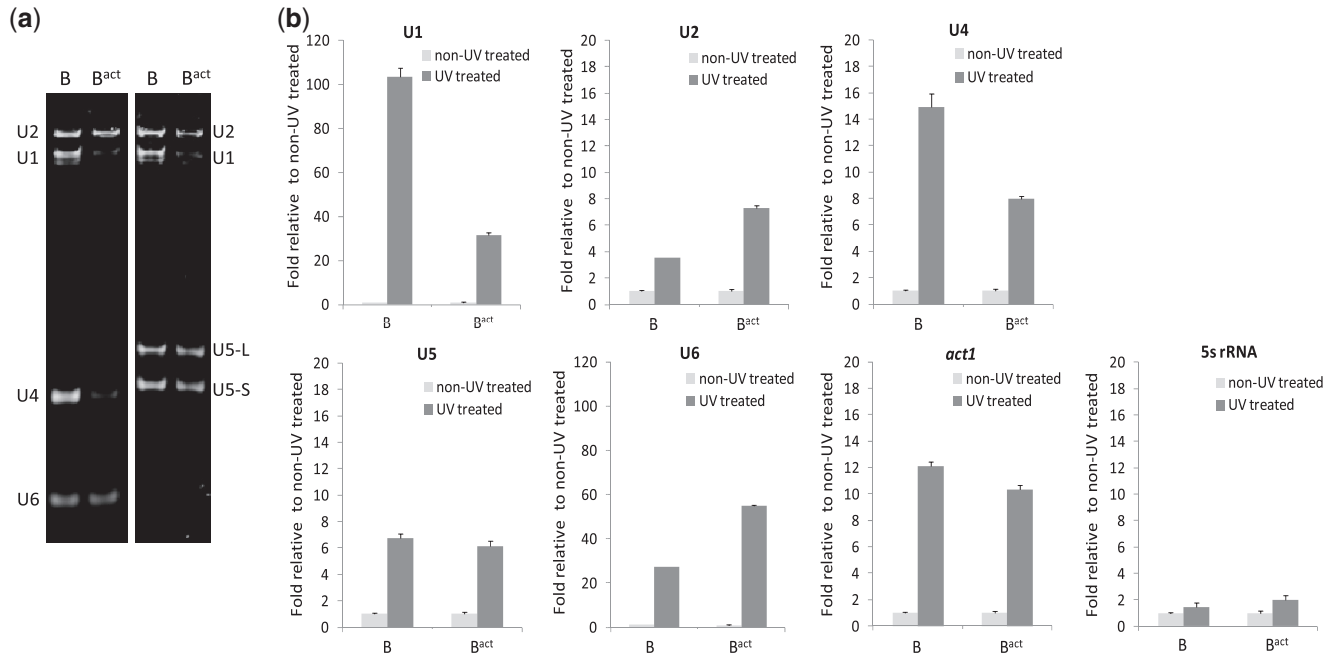


Figure 6. Prp8 and snRNAs interactions in purified spliceosomal B and B^{act} complex. (**a**) RNAs extracted from purified spliceosomal B and B^{act} complexes, hybridized with fluorescently labelled probes specific for U1, U2, U4, U5 and U6 snRNAs in solution, and analyzed on a native polyacrylamide gel, demonstrated that the B^{act} complex has significantly lowered U1 and U4 snRNA levels compared with the B complex. (**b**) Real-time PCR quantification of snRNAs associated with Prp8 in B and B^{act} complex after UV cross-linking. All samples are normalized to the non–UV-treated sample. 5S rRNA is used as negative controls.

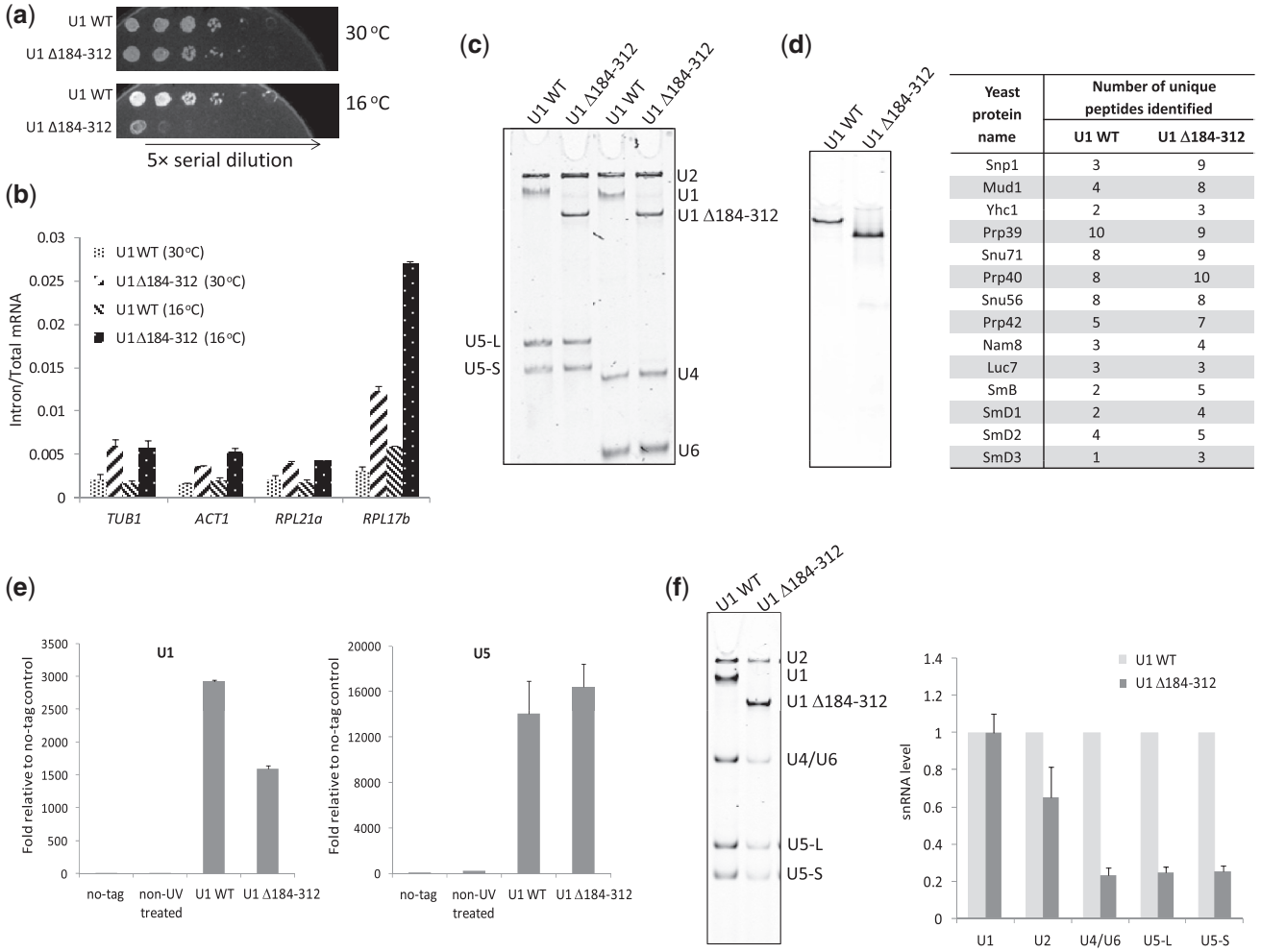
real-time PCR. These experiments demonstrate that U1 Δ184–312 has ∼2-fold reduction in its cross-linking with Prp8, whereas U5 snRNA cross-links with Prp8 at similar levels in both the U1 Δ184–312 and WT strains (Figure 7e).

We carried out a spliceosomal assembly assay using the yeast M3-Act1 Δ6 pre-mRNA substrate incubated with splicing extract at 0.05 mM ATP, which leads to the formation of the spliceosomal B complex (23). After affinity purification of the assembled spliceosome, we analysed RNAs on a native polyacrylamide gel using in solution hybridization (35) with fluorescent probes specific to each snRNA (Figure 7f). These analyses demonstrate that the assembled spliceosome in the U1 Δ184–312 strain has similar levels of U1 snRNAs but significantly reduced levels of U4, U5 and U6 snRNAs.

DISCUSSION

We have determined the comprehensive *in vivo* RNA footprints of Prp8 in budding yeast using CLIP/CRAC methods. Prp8 predominantly binds snRNAs and intronic pre-mRNAs. The binding sites of Prp8 on U5, U6 and pre-mRNA have been previously investigated using 4-thiouridine incorporated at specific positions of the RNA with *in vitro* assembled U5, tri-snRNP and the spliceosome. Prp8 was found to cross-link to 4-thiouridine incorporated at positions 20, 59, 97, 112 and 134 in reconstituted yeast U5 snRNA (28), position 54 in U6 snRNA in reconstituted tri-snRNPs (11) and regions spanning positions −8 to +3 relative to the 5′ ss, around the BPS, and from positions −12 to +13 relative to the 3′ ss [reviewed in (3)]. All of these *in vitro* cross-linking sites are within the *in vivo* footprints of Prp8 we identified using CLIP/CRAC. U5 snRNA is the most abundantly detected snRNA in the sequencing reads, and Prp8 binds predominantly to a highly protected ∼75-nt region (positions 59–130) including loop 1. The cross-linking between Prp8 and U5 loop 1 is also observed in human tri-snRNP purified from HeLa cells (36), demonstrating striking evolutionary conservation. The large footprint of Prp8 on U5 snRNA observed in our CLIP/CRAC experiments explains why deletion of loop 1 positions 93–101 of U5 snRNA does not significantly affect its co-precipitation with Prp8, but its cross-linking with Prp8 is reduced by ∼50% (28). U6 is the second most abundant snRNA detected followed by U2 and U1 snRNAs. The Prp8 footprints on these snRNAs reflect a combination of binding sites that could come from individual snRNP, U4/U6 di-snRNP, tri-snRNP or the spliceosomal complex.

A significant number of sequencing reads map to positions 32–70 in U4 snRNA in CLIP experiments but not in CRAC experiments. The sequencing reads could come from another spliceosomal protein, which is tightly associated with Prp8 and not completely removed in CLIP experiments. Brr2 is a potential candidate because of its tight association with Prp8 (37) and its similar size. Indeed, a recent CRAC analyses indicate that Brr2 binds positions 30–75 in U4 snRNA (38). Alternatively, these sequencing reads may be a result of the base pairing in stem 1, which is not completely disrupted in CLIP experiments. Kudla *et al.* observed undisrupted RNA base pairing even after nickel purification under 6 M



Yeast protein name	Number of unique peptides identified	
	U1 WT	U1 Δ184-312
Snp1	3	9
Mud1	4	8
Yhc1	2	3
Prp39	10	9
Snu71	8	9
Prp40	8	10
Snu56	8	8
Prp42	5	7
Nam8	3	4
Luc7	3	3
SmB	2	5
SmD1	2	4
SmD2	4	5
SmD3	1	3

Figure 7. The function of the interaction between Prp8 and U1 snRNA. (**a**) U1 Δ184–312 has a cold-sensitive growth phenotype. (**b**) U1 Δ184–312 strain has splicing defects at both 30°C and 16°C, manifested as pre-mRNA accumulation in a number of genes we examined using real-time PCR. (**c**) U1 Δ184–312 does not affect U1 and other snRNA levels in the cell when total cellular RNA is quantified using in solution hybridization with probes specific to each snRNA. (**d**) U1 snRNP purified by affinity pull-down of U1-70K has similar U1 snRNA (left, by in solution hybridization) and protein components (right, by mass spectrometry). Sm E, F and G proteins are not identified in the U1 snRNP from either the WT or U1 Δ184–312 strain. (**e**) U1 Δ184–312 has reduced cross-linking to Prp8 by ~2-fold compared with the WT, whereas U5 snRNA cross-links to Prp8 at similar levels in both strains. (**f**) Spliceosomal B complex assembled from the U1 Δ184–312 strain has similar U1 snRNA level but much reduced U4, U5 and U6 levels compared with the WT. The gel bands are quantified on the right with snRNA levels in the WT normalized to 1.

guanidine–HCl in CRAC experiments (30). In either case, these data suggest that positions 32–70 in U4 snRNA (half of the 5′ stem–loop and stem 1) are well protected from RNase digestion by proteins in the cell. The Brr2-binding sites on U5 and U6 snRNAs from the CRAC analyses (38) overlap with the Prp8-binding sites on these snRNAs, although the number of sequencing reads of Brr2 at these sites is much lower than that of Prp8. It is possible that Prp8 and Brr2 bind to these sites at different stages of the splicing process. Alternatively, Prp8 and Brr2 may bind to different sides of the RNA duplex at the same time, resulting in overlapping RNA footprints.

Analyses of deletions in CRAC sequencing reads can reveal direct cross-linking sites on RNAs. These analyses revealed direct cross-link sites of Prp8 on loop1 (96–99 nt) of U5 snRNA and the 5′ ss +1 position of pre-mRNA, consistent with previously observed *in vitro* cross-linking sites. These analyses also reveal additional Prp8 direct cross-linking sites at positions 44–45 of U6 (in addition to U54 examined with *in vitro* cross-linking experiments) and positions 16–19 of U2 snRNAs. We note that the direct cross-linking sites revealed by CRAC are not exhaustive. The amino acids and nucleotide bases must be at the appropriate distance and have the distinct chemical properties to be cross-linked by UV radiation at 254 nm (14,15). Although the CRAC experiments may not identify each nucleotide in the footprint that directly interacts with Prp8, these experiments can identify comprehensive *in vivo* RNA footprints of Prp8 that are difficult to obtain using other methods.

In addition to mapping the composite Prp8 RNA footprints, we examined Prp8 footprints in purified U5, tri-snRNP, spliceosomal B and B^{act} complexes. The Prp8 footprints on U5 snRNA are similar in U5 and tri-snRNP, suggesting that Prp8 does not significantly change its overall binding sites, but it gains additional U4- and U6-

binding sites (likely through different parts of the Prp8) in tri-snRNP compared with U5 snRNP. Prp8 footprints are enriched at the 3′ half of U4 snRNA, which are not obvious in the whole-cell CRAC data set. This likely reflects a weak Prp8-binding site on U4 snRNA, which is more abundant in the tri-snRNP data set, in which U4 snRNA is greatly enriched, and the RNase treatment dose is much lower compared with the whole-cell data set. Prp8 seems to contact all snRNAs and pre-mRNA in the spliceosomal B complex (Figure 6). Its contacts with U5 snRNA and pre-mRNA remain similar in the B and B^{act} complexes, but its contacts with U1 and U4 snRNA are significantly reduced, consistent with the release of U1 and U4 snRNAs upon the formation of the B^{act} complex. The interaction between Prp8 and U1 and U4 snRNAs do not completely disappear, likely because the B^{act} complex assembled and purified under this condition often contains residual U1 and U4 snRNAs (Figure 6a). Interestingly, Prp8 seems to be already in contact with U2 and U6 snRNAs in the B complex, and these contacts increased from the B to the B^{act} complex. These data provide a dynamic view of how Prp8-binding sites evolve in different snRNP and spliceosomal complexes, suggesting a potential role of Prp8 in coordinating and assisting the transition between different snRNPs and spliceosomal complexes.

Mapping the Prp8 footprints and cross-linking sites onto the proposed spliceosome active site configuration reveal interesting insight into the relationship between Prp8 and the spliceosome active site (Figure 8). In the proposed spliceosome active site configuration, the evolutionarily invariant ACAGAGA box of U6 interacts with the 5′ ss, and the 5′-end of U2 interacts with the BPS. The same regions of U2 and U6 also base pair with each other, bringing the branchpoint adenosine and the 5′ ss close to each other. U5 interacts with both exons, which is likely important for aligning them during the second step catalytic reaction [reviewed in (40,41)]. Most of these interactions (with the exception of U2–U6 base pairing) are between very short RNA stretches, which are unlikely to be stable on their own. On the other hand, the footprints of Prp8 centre at positions 59–130 (including loop 1) in U5 snRNA, positions 44–70 (including the ACAGAGA and AGC boxes) in U6 snRNA, position 3–67 in U2 snRNA, the 5′ ss, 3′ ss and BPS. Prp8 cross-links to loop 1 of U5, A44–A45 and U54 of U6 (immediately upstream and downstream of ACAGAGA), nucleotides 16–19 of U2 snRNA, as well as the 5′ ss, BPS and 3′ ss. These footprints and cross-linking sites put Prp8 in a perfect position to be the protein factor that brings all components of the active site together, to help form and stabilize the catalytic core. One caveat of the CRAC experiments is that these experiments provide a composite RNA-binding map of Prp8 in the entire yeast cell, and we do not know during which stage and complex Prp8 binds to a particular site. As a first step to address this issue, we showed that Prp8 cross-links with U2, U5, U6 snRNAs, and the *act1* pre-mRNA in purified spliceosomal B^{act} complex. A potential role of Prp8 in helping to form and stabilize the catalytic core has been suggested based on previous genetic data and *in vitro* cross-linking experiments with isolated

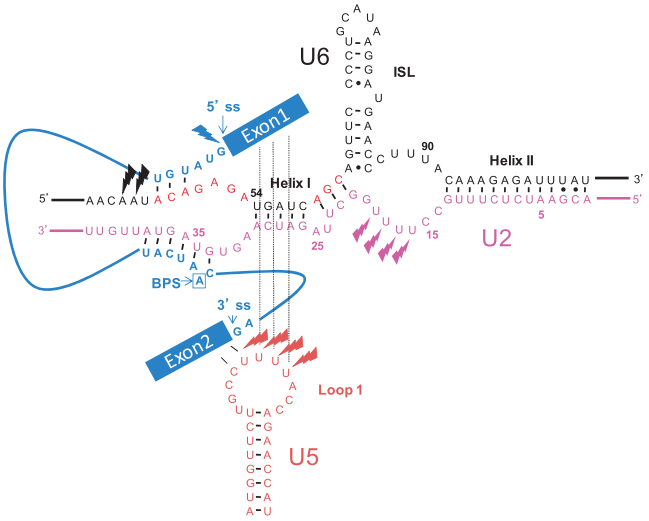


Figure 8. A schematic representation of the proposed spliceosome active site conformation [modified from (39)]. U5, U2 and U6 snRNAs and the pre-mRNA are coloured in brown, purple, black and dark blue, respectively. The invariant ACAGAGA and AGC boxes in U6 are shown in red. Bolts represent major cross-linking sites observed in our CLIP/CRAC experiments.

snRNPs [(28,42) and reviewed in (3)]. The *in vivo* RNA footprints and cross-linking sites of Prp8 identified by CRAC, as well as our observation that Prp8 simultaneously and directly cross-links to U2, U5, U6 and pre-mRNA in purified spliceosomal B^{act} complex (Figure 6), provide further support for this hypothesis.

Our CRAC analyses also revealed surprising Prp8-binding sites on U1 snRNA. Yeast U1 snRNA is 3.5× larger than its human counterpart (568 versus 164 nt) (32). There are clear sequence homologies between the yeast and human U1 snRNA in the 5′ ss base pairing region the long range interaction region (LRI), stem/loop 1 that binds U1-70K and the Sm-binding site (Figure 3a). The remaining sequences have no obvious sequence homology, and nucleotides 46–167 were tentatively assigned stem/loop 2 and the remaining nucleotides as stem/loop 3 (32). With this assignment, the Prp8-binding site mostly falls on stem/loop 3, whose human counterpart is not bound by any U1 snRNP proteins in the crystal structure (43), making it a feasible position for Prp8 binding. Prp8 is likely to bind stem/loop 3 as well in human U1 snRNA, but its binding site is likely not as extensive as in yeast U1 snRNA, as the stem/loop 3 in human U1 snRNA is much shorter than in yeast.

To understand the function of the Prp8 and U1 snRNA interaction, we used a U1 deletion construct (U1 Δ184–312) with partial Prp8-binding sites deleted that has reduced Prp8 cross-linking and cold-sensitive growth, but it does not affect the overall U1 snRNA level in the cell or U1 snRNP assembly. We found that spliceosomal B complex assembled using yeast extract from the U1 Δ184–312 strain has similar U1 snRNA level, but it had significantly reduced U4, U5 and U6 snRNA levels compared with the WT control. Our results suggest that the interaction between Prp8 and U1 snRNA is

important for recruiting the tri-snRNP during spliceosomal assembly. The U2 snRNA level in B complex from the U1 Δ184–312 strain is also slightly lower than the WT, suggesting that U2 snRNP association with the spliceosome is unstable without the engagement of tri-snRNP. It was previously unclear what recruits the tri-snRNP to the spliceosome to form the B complex. Our results provide one answer for this missing piece in the spliceosomal assembly process.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online: Supplementary Methods and Supplementary Reference [44].

ACKNOWLEDGEMENTS

The authors thank Dr Tom Blumenthal at the University of Colorado Boulder for critical reading of the manuscript, Dr Mark Johnston and Jim Dover in the Department of Biochemistry and Molecular Genetics, Jill Castoe at the high-throughput sequencing core at the University of Colorado School of Medicine and David Mohr at the high-throughput sequencing centre at Johns Hopkins University for help with high-throughput sequencing.

FUNDING

National Institutes of Health grant [R01GM080334 to R. Z.]; American Heart Association [postdoctoral fellowships to T.X. and L.Z.]; Damon Runyon Cancer Research Foundation [Damon Runyon-Rachleff Innovation Award (DRR-17-12) and the March of Dimes (5-FY10-478-02) to J.R.H.]. Funding for open access charge: NIH [R01GM080334].

Conflict of interest statement. None declared.

REFERENCES

1. Moore,M.J., Query,C.C. and Sharp,P.A. (1993) Splicing of precursors to mRNA by the spliceosome. In: Gesteland,R. and Atkins,J. (eds), *The RNA World*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp. 303–357.
2. Jurica,M.S. and Moore,M.J. (2003) Pre-mRNA splicing: awash in a sea of proteins. *Mol. Cell*, **12**, 5–14.
3. Grainger,R.J. and Beggs,J.D. (2005) Prp8 protein: at the heart of the spliceosome. *RNA*, **11**, 533–557.
4. Pena,V., Liu,S., Bujnicki,J.M., Luhrmann,R. and Wahl,M.C. (2007) Structure of a multipartite protein-protein interaction domain in splicing factor prp8 and its link to retinitis pigmentosa. *Mol. Cell*, **25**, 615–624.
5. Zhang,L., Shen,J., Guarnieri,M.T., Heroux,A., Yang,K. and Zhao,R. (2007) Crystal structure of the C-terminal domain of splicing factor Prp8 carrying retinitis pigmentosa mutants. *Protein Sci.*, **16**, 1024–1031.
6. Yang,K., Zhang,L., Xu,T., Heroux,A. and Zhao,R. (2008) Crystal structure of the beta-finger domain of Prp8 reveals analogy to ribosomal proteins. *Proc. Natl Acad. Sci. USA*, **105**, 13817–13822.
7. Pena,V., Rozov,A., Fabrizio,P., Luhrmann,R. and Wahl,M.C. (2008) Structure and function of an RNase H domain at the heart of the spliceosome. *EMBO J.*, **27**, 2929–2940.
8. Ritchie,D.B., Schellenberg,M.J., Gesner,E.M., Raithatha,S.A., Stuart,D.T. and Macmillan,A.M. (2008) Structural elucidation of a PRP8 core domain from the heart of the spliceosome. *Nat. Struct. Mol. Biol.*, **15**, 1199–1205.
9. Kuhn,A.N. and Brow,D.A. (2000) Suppressors of a cold-sensitive mutation in yeast U4 RNA define five domains in the splicing factor Prp8 that influence spliceosome activation. *Genetics*, **155**, 1667–1682.
10. Kuhn,A.N., Reichl,E.M. and Brow,D.A. (2002) Distinct domains of splicing factor Prp8 mediate different aspects of spliceosome activation. *Proc. Natl Acad. Sci. USA*, **99**, 9145–9149.
11. Vidal,V.P., Verdone,L., Mayes,A.E. and Beggs,J.D. (1999) Characterization of U6 snRNA-protein interactions. *RNA*, **5**, 1470–1481.
12. Ule,J., Jensen,K., Mele,A. and Darnell,R.B. (2005) CLIP: a method for identifying protein-RNA interaction sites in living cells. *Methods*, **37**, 376–386.
13. Granneman,S., Kudla,G., Petfalski,E. and Tollervey,D. (2009) Identification of protein binding sites on U3 snoRNA and pre-rRNA by UV cross-linking and high-throughput analysis of cDNAs. *Proc. Natl Acad. Sci. USA*, **106**, 9613–9618.
14. Brimacombe,R., Stiege,W., Kyriatsoulis,A. and Maly,P. (1988) Intra-RNA and RNA-protein cross-linking techniques in *Escherichia coli* ribosomes. *Methods Enzymol.*, **164**, 287–309.
15. Shetlar,M.D., Christensen,J. and Hom,K. (1984) Photochemical addition of amino acids and peptides to DNA. *Photochem. Photobiol.*, **39**, 125–133.
16. Umen,J.G. and Guthrie,C. (1996) Mutagenesis of the yeast gene PRP8 reveals domains governing the specificity and fidelity of 3′ splice site selection. *Genetics*, **143**, 723–739.
17. Ghaemmaghami,S., Huh,W.K., Bower,K., Howson,R.W., Belle,A., Dephoure,N., O'Shea,E.K. and Weissman,J.S. (2003) Global analysis of protein expression in yeast. *Nature*, **425**, 737–741.
18. Longtine,M.S., McKenzie,A. 3rd, Demarini,D.J., Shah,N.G., Wach,A., Brachat,A., Philippsen,P. and Pringle,J.R. (1998) Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast*, **14**, 953–961.
19. Grate,L. and Ares,M. Jr (2002) Searching yeast intron data at ares lab web site. *Methods Enzymol.*, **350**, 380–392.
20. Christie,K.R., Weng,S., Balakrishnan,R., Costanzo,M.C., Dolinski,K., Dwight,S.S., Engel,S.R., Feierbach,B., Fisk,D.G., Hirschman,J.E. *et al.* (2004) *Saccharomyces* genome database (SGD) provides tools to identify and analyze sequences from *Saccharomyces cerevisiae* and related sequences from other organisms. *Nucleic Acids Res.*, **32**, D311–D314.
21. Pleiss,J.A., Whitworth,G.B., Bergkessel,M. and Guthrie,C. (2007) Transcript specificity in yeast pre-mRNA splicing revealed by mutations in core spliceosomal components. *PLoS Biol.*, **5**, e90.
22. Stevens,S.W., Barta,I., Ge,H.Y., Moore,R.E., Young,M.K., Lee,T.D. and Abelson,J. (2001) Biochemical and genetic analyses of the U5, U6, and U4/U6 x U5 small nuclear ribonucleoproteins from *Saccharomyces cerevisiae*. *RNA*, **7**, 1543–1553.
23. Fabrizio,P., Dannenberg,J., Dube,P., Kastner,B., Stark,H., Urlaub,H. and Luhrmann,R. (2009) The evolutionarily conserved core design of the catalytic activation step of the yeast spliceosome. *Mol. Cell*, **36**, 593–608.
24. McManus,C.J., Schwartz,M.L., Butcher,S.E. and Brow,D.A. (2007) A dynamic bulge in the U6 RNA internal stem-loop functions in spliceosome assembly and activation. *RNA*, **13**, 2252–2265.
25. Lu,X. and Zhu,H. (2005) Tube-gel digestion: a novel proteomic approach for high throughput analysis of membrane proteins. *Mol. Cell. Proteomics*, **4**, 1948–1958.
26. O'Brien,J.H., Vanderlinden,L.A., Schedin,P.J. and Hansen,K.C. (2012) Rat mammary extracellular matrix composition and response to ibuprofen treatment during postpartum involution by differential GeLC-MS/MS analysis. *J. Proteome Res.*, **11**, 4894–4905.
27. Frank,D.N., Roiha,H. and Guthrie,C. (1994) Architecture of the U5 small nuclear RNA. *Mol. Cell Biol.*, **14**, 2180–2190.

28. Dix,I., Russell,C.S., O'Keefe,R.T., Newman,A.J. and Beggs,J.D. (1998) Protein-RNA interactions in the U5 snRNP of *Saccharomyces cerevisiae*. *RNA*, **4**, 1675–1686.
29. Zhang,C. and Darnell,R.B. (2011) Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. *Nat. Biotechnol.*, **29**, 607–614.
30. Kudla,G., Granneman,S., Hahn,D., Beggs,J.D. and Tollervey,D. (2011) Cross-linking, ligation, and sequencing of hybrids reveals RNA-RNA interactions in yeast. *Proc. Natl Acad. Sci. USA*, **108**, 10010–10015.
31. Madhani,H.D., Bordonne,R. and Guthrie,C. (1990) Multiple roles for U6 snRNA in the splicing pathway. *Genes Dev.*, **4**, 2264–2277.
32. Siliciano,P.G., Kivens,W.J. and Guthrie,C. (1991) More than half of yeast U1 snRNA is dispensable for growth. *Nucleic Acids Res.*, **19**, 6367–6372.
33. Shuster,E.O. and Guthrie,C. (1988) Two conserved domains of yeast U2 snRNA are separated by 945 nonessential nucleotides. *Cell*, **55**, 41–48.
34. Yan,D. and Ares,M. Jr (1996) Invariant U2 RNA sequences bordering the branchpoint recognition region are essential for interaction with yeast SF3a and SF3b subunits. *Mol. Cell Biol.*, **16**, 818–828.
35. Li,Z. and Brow,D.A. (1993) A rapid assay for quantitative detection of specific RNAs. *Nucleic Acids Res.*, **21**, 4645–4646.
36. Urlaub,H., Hartmuth,K., Kostka,S., Grelle,G. and Luhrmann,R. (2000) A general approach for identification of RNA-protein cross-linking sites within native human spliceosomal small nuclear ribonucleoproteins (snRNPs). Analysis of RNA-protein contacts in native U1 and U4/U6.U5 snRNPs. *J. Biol. Chem.*, **275**, 41458–41468.
37. Achsel,T., Ahrens,K., Brahms,H., Teigelkamp,S. and Luhrmann,R. (1998) The human U5-220kD protein (hPrp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor EF-2, and a novel WD-40 protein. *Mol. Cell Biol.*, **18**, 6756–6766.
38. Hahn,D., Kudla,G., Tollervey,D. and Beggs,J.D. (2012) Brr2p-mediated conformational rearrangements in the spliceosome during activation and substrate repositioning. *Genes Dev.*, **26**, 2408–2421.
39. Valadkhan,S. (2010) Role of the snRNAs in spliceosomal active site. *RNA Biol.*, **7**, 345–353.
40. Valadkhan,S. (2005) snRNAs as the catalysts of pre-mRNA splicing. *Curr. Opin. Chem. Biol.*, **9**, 603–608.
41. Collins,C.A. and Guthrie,C. (2000) The question remains: is the spliceosome a ribozyme? *Nat. Struct. Biol.*, **7**, 850–854.
42. Collins,C.A. and Guthrie,C. (1999) Allele-specific genetic interactions between Prp8 and RNA active site residues suggest a function for Prp8 at the catalytic core of the spliceosome. *Genes Dev.*, **13**, 1970–1982.
43. Pomeranz Krummel,D.A., Oubridge,C., Leung,A.K., Li,J. and Nagai,K. (2009) Crystal structure of human spliceosomal U1 snRNP at 5.5 a resolution. *Nature*, **458**, 475–480.
44. Li,Z. and Brow,D.A. (1996) A spontaneous duplication in U6 spliceosomal RNA uncouples the early and late functions of the ACAGA element *in vivo*. *RNA*, **2**, 879–894.